

**CHARACTERIZATION OF PHENOTYPES BY GENE EXPRESSION  
PATTERNS AND CLASSIFICATION OF SAMPLES BASED THEREON**

**Abstract of the Disclosure**

5

Generally, the present invention applies a transformation to convert a probability distribution of gene expression signals in control samples to a uniform distribution. The uniform distribution allows better comparisons between expression levels for genes. The transformation is derived from gene expression signals of control 10 data, and is applied to gene expression signals of phenotype data. The phenotype data can be represented in a matrix format. A number of gene expression patterns may be determined (in the form of submatrices) that will characterize the phenotype. The uniform distribution helps in this regard, as it allows better comparisons of patterns. The gene expression patterns can then be used to classify samples as belonging to the phenotype 15 set. Preferably, a discriminant function is used to compare a sample with the gene expression patterns that characterize the phenotype. The discriminant function can determine a score that can be used to determine whether the sample belongs to the phenotype.